2010

An Evaluation of a Dialogic Book-Reading Program for At Risk Children

Daniel Anthony Colangelo
Wilfrid Laurier University

Your file Votre référence
ISBN: 978-0-494-68719-2
Our file Notre référence
ISBN: 978-0-494-68719-2

AN EVALUATION OF A DIALOGIC BOOK-READING PROGRAM FOR AT RISK CHILDREN

By

Daniel Anthony Colangelo

Bachelor of Arts, York University, 2007

Master's Thesis

Submitted to the Department of Psychology

in partial fulfillment of the requirements for the Master of Arts

Wilfrid Laurier University

2010

ABSTRACT

Children from low-income backgrounds are at a higher risk for reading difficulties partly because they are read to less frequently in the home (Adams, 1990). When shared reading does occur in low-income homes, it is usually of poorer quality when compared to reading in middle- or upper-income homes (Arnold, Lonigan, Whitehurst, Epstein, 1994). Dialogic reading, a form of enhanced discussion and structured questioning during shared-book reading, can be a cost effective way of improving the language and literacy skills of young children. The current research examines the effectiveness of a community-based, four-month dialogic reading intervention called the Dialogic Reading Club (pseudonym used to protect the identity of the program). Eighteen children aged 38 months to 68 months (*M* age = 58.22 months) that attended the intervention were compared with 18 children aged 39 months to 71 months (*M* age = 53.11 months) that did not attend the intervention on measures of expressive vocabulary, word reading, concepts of print, and narrative ability. Controlling for pre-test differences on the same post-test measures, ANCOVAs revealed significant differences in word reading, $F(1,33) = 5.40$, $p<.05$, and a measure of concepts of print, $F(1,33) = 9.28$, $p<.05$ in favour of the intervention group. Differences in narrative structure and ability approached significance, with ANOVAs revealing that children in the intervention group produced higher quality narratives ($p = .09$), produced more words ($p = .08$), and produced a greater diversity of words ($p = .08$). No differences were found on expressive vocabulary. The benefits of incorporating dialogic reading strategies in a short-term reading intervention for young children are discussed.

ACKNOWLEDGEMENTS

The hard word and dedication of many people allowed this project to come to fruition. Firstly, I would like to thank the coordinator of the Dialogic Reading Club for her assistance throughout the study. Secondly, I would like to thank all the participants. Their hard work and patience through two lengthy assessments is greatly appreciated.

I owe a great deal of gratitude towards the team of researchers for their assistance with data collection. Their time and effort allowed this project to be completed in a timely manner. I enjoyed working as part of this dedicated team of researchers.

I wish to thank a number of people for their support during my time at Laurier. I thank Amy Grant for her continuous encouragement and guidance. I thank Fanli Jia for his friendship and genorisity. I thank Dr. Gottardo for imparting to me her extensive knowledge in the field of reading, for always offering useful suggestions, and for offering continuous positive support and care. You have inspired me as a researcher and as an individual. I look forward to the future and to continue to collaborate on research.

Finally, I thank my family. Mom, Dad, John, and Anthony – I owe the greatest thanks to you for your love and support.

TABLE OF CONTENTS

LIST OF TABLES

An Evaluation of a Dialogic Book-Reading Program for At Risk Children

Despite the proven positive effects of dialogic reading, there has been a dearth of studies in recent years that further examine this form of reading. Much of the recent research in the field of reading interventions has focused on phonological awareness interventions, which teach letter-sound matching and how to manipulate parts of words. While phonological awareness interventions can help young children improve their word reading ability, they do little to improve oral language skills and other emergent literacy skills. Dialogic reading, characterized by increased dialogue during shared storybook reading, can be a cost-effective method of improving the literacy skills of children, specifically expressive vocabulary and oral language skills. Dialogic reading entails the child becoming the storyteller, with an adult taking the role of active listener (Arnold, Lonigan, Whitehurst, & Epstein, 1994). Dialogic reading provides children the opportunity to express themselves with the aid of structured questions, thereby building oral language skills. Since dialogic reading targets a variety of important literacy skills besides word reading, it is important to analyze this form of intervention further.

This study focuses on a dialogic reading intervention for children from low-income homes in Toronto, Ontario. This study has the applied goal of improving the literacy skills of this specific group of at risk children. While the program coordinator has noted the effectiveness of dialogic reading in improving children's literacy skills, this study will serve to validate these claims scientifically. In terms of research, this study will add to the literature on dialogic reading by using a wider array of assessment instruments, using a Canadian sample, and utilizing community volunteers instead of parents or teachers, which are typically the focus of dialogic reading intervention studies. Using a Canadian sample is important as Canada offers a diverse population. There is a large proportion of recent immigrants in the communities in the

present study, which provides the unique opportunity to see the effects of dialogic reading for children whose parents have English as a second language (ESL).

The present study focuses on children that attend community centers that serve generally low-income families. The Toronto Neighborhood Profile website (Toronto.ca/demographics) provides useful demographic information for the communities in the present study. One community center is located in a community that serves a population with a higher percentage of immigrants compared with the rest of Toronto and where 26.5% of families are classified as low-income. In the second community where the program is offered, 26.9% of families are classified as low-income. Families categorized as at risk either receive government support, have low incomes, or have less educated mothers (Mol, Bus, de Jong, & Smeets, 2008). This present study, then, focuses on at risk children. Additionally the ESL status of the children in these communities makes them at risk for oral language delays in comparison to their English-speaking peers (August, Carlo, Dressler & Snow, 2005).

In order to put this study in context, it is necessary to examine the relevant literature. To begin, the discrepancy in literacy skills between children from low-income families and children from middle- and upper-income families will be discussed. A large body of research indicates that a reason for this discrepancy stems from differences in shared storybook reading, both in quality and quantity (Adams, 1990; Snow, Burns, & Griffin, 1998). Different ways to enhance storybook reading will be followed by a thorough definition of the concept of dialogic reading, an established strategy used to enhance storybook reading. Various interventions using dialogic reading will be analyzed. Finally, the effectiveness of using appropriately levelled print will be discussed, as this is a key component of the current intervention.

Literacy for low-income children

In a comprehensive review of literacy among children, Adams (1990) claimed that children from low-income backgrounds are at higher risk for reading failure and begin school less prepared to read than their more affluent peers. The suggestion that socioeconomic status is strongly related to literacy and school success has been made by numerous other researchers (Payne, Whitehurst, & Angell, 1994; Snow, Burns, & Griffin, 1998; Whitehurst & Lonigan, 1998; Arnold, Lonigan, Whitehurst, & Epstein, 1994). Children from poor families and children attending urban schools are at much greater risk of poor reading outcomes and lower overall academic achievement than are middle-class and suburban children (Snow, Burns, & Griffin, 1998).

Researchers have speculated many potential reasons for the discrepancy in literacy level between children from low-income backgrounds and children from middle- or upper-class backgrounds. Harris and Smith (1987) claimed that children from low-income families are read to less frequently than children from higher socio-economic status (SES) groups. Since shared storybook reading is an important element in the development of literacy skills (Adams, 1990; Scarborough & Dobrich, 1994), a lack of this experience may impede literacy development for low-income children. For preschool children, shared storybook reading typically takes the form of a parent reading a book to their child. Adams (1990, p.85) estimated that a typical middle-class child enters first grade with 1,000 to 1,700 hours of one-on-one picture book reading, whereas a child from a low-income family averages just 25 such hours. Whitehurst and Lonigan (1998) also believe that the social class differences that exist in oral language and pre-literacy skills are associated with the large discrepancy in the amount of shared-reading in the home.

When shared reading occurs in low-income homes, it is usually of poorer quality when compared to reading in middle- or upper-income homes. Several researchers (Arnold et al., 1994; Ninio, 1980) found that parents of children from lower SES groups typically engaged in fewer instructive behaviors during story reading. For example, mothers from low-SES groups were less likely to label object attributes and actions, and were less likely to ask "where" or "what" questions when reading with their child. Labeling object attributes and actions and asking questions during storybook reading are helpful for the acquisition of vocabulary (Senechal, 1995).

Many researchers have examined the link between home literacy environment and child language ability. Payne et al. (1994) examined this relationship in 323 four-year-olds attending Head Start programs, an organization that serves low-income families. A composite literacy environment score was calculated. The home literacy environment score was derived from measures of frequency of shared picture book reading; age of onset of picture book reading; duration of shared picture book reading during one recent day; number of picture books in the home; frequency of child's requests to engage in shared picture book reading; frequency of child's private play with books; frequency of shared trips to the library; frequency of caregiver's private reading; and caregiver's enjoyment of private reading and correlated with a composite child language measure using two standardized tests of language skills, one that measured receptive vocabulary and another that measured expressive vocabulary. When primary caregiver IQ and education were entered into the prediction equations, 18.5% of the variance in child language scores was accounted for by home literacy environment (12.5% when primary caregiver IQ and education were excluded from the model). This study demonstrated that there are significant differences among low-income families in home literacy environment and that

these differences are associated with children's language development. Large differences in home literacy environments can be found within low-SES groups. Between SES groups, this difference is even larger (Adams, 1990).

Several other studies also link aspects of children's home environment, parent demographics, and SES to reading achievement. There is evidence indicating that children's vocabulary sizes are correlated with parental education and indicators of environmental quality (Hall, Nagy, & Linn, 1984). McCormick and Mason (1986) demonstrated large social class differences in the availability of printed material in the home. Children whose home literacy environments are lacking in terms of shared reading activities and print materials are also likely to have poor oral language skills (Storch & Whitehurst, 2002).

Children from low-income backgrounds are particularly likely to have low levels of oral language and emergent literacy skills, which are important in formal schooling, and such children are at risk for later reading difficulties (Raz & Bryant, 1990). Children from low-income families are likely to start school behind and stay behind (Whitehurst & Lonigan, 1998). This late start in literacy has far-reaching consequences. Early success at reading acquisition can help develop a lifetime of reading habits. A longitudinal study by Stanovich and Cunningham (1997) tested 27 eleventh-grade students, ten years after they had been initially tested in first-grade. Measures of first grade reading ability, word reading and reading comprehension, predicted significant variance in eleventh-grade print exposure (magazine recognition checklist and author recognition checklist), even after eleventh-grade reading comprehension ability had been partialed out. This study demonstrated that an early start in reading is important in predicting a lifetime of literacy experience. If a child gets off to a fast start in reading, they are

more likely to engage in more reading activity, even when taking into account the subsequent level of reading comprehension ability.

Print exposure and storybook reading

Preschool children acquire literacy knowledge and skills that correlate with later literacy development and academic success, including oral language, knowledge about print, phonemic awareness, and understanding about oral and written language and their distinctions (Arnold, Lonigan, et al., 1994). Shared picture book reading is thought to provide an important environment for the acquisition of these preliteracy skills (Adams, 1990). Research conducted by Cunningham and Stanovich (1997) suggested that exposure to print serves to develop processes and knowledge bases, including vocabulary and familiarity with complex syntactic structures that facilitate reading comprehension. Cunningham and Stanovich (1997) believe that "one of the most powerful determinants of individual differences in vocabulary is exposure to print" (p. 265).

Exposure to books can have wide-ranging impacts on a variety of reading-related skills. Research suggests that print exposure is a critical source for the development of vocabulary (Mol, Bus, de Jong, & Smeets, 2008). Senechal and LeFevre (2001) claim that "children learn vocabulary from listening to adults read books to them" (p. 41). Greater print exposure can result in more experience in word decoding, sentence decoding, reading comprehension, and ongoing exposure to new vocabulary (Hood, Conlon, & Andrews, 2008). Amount of reading experience has a large impact on comprehension ability, which may result from an increased familiarity with narrative structures (Cain, 1996). Print exposure also contributes significantly

to the code-related skill base of emergent literacy, including knowledge of the conventions of print (e.g., knowing that writing goes from left to right).

A study by Elley (1989) of seven- and eight-year-old New Zealand children provided further evidence that children can learn new vocabulary incidentally from having illustrated storybooks read to them. During a pre-test, children were given four options and asked to choose the best definition (half the items) or picture for 20 target words. The following week, teachers read children a picture book containing the words from the pretest. One week later, children took the same vocabulary test. Overall, children scored higher on most target words on the post-test than on the pre-test, for a mean increase of 15.4% overall. This effect was replicated for two different picture books, with similar effect sizes. The number of times the book was read over the course of one week, ranging from one to three readings, was positively correlated with the amount of vocabulary learned. Also, children in a select group who were exposed to teachers that provided brief explanations of unknown words as they were encountered in text doubled their vocabulary gains. Although long-term vocabulary gains can be questioned in this study, oral story reading constitutes a significant source of immediate vocabulary acquisition, and even more so when brief explanations of unknown words are provided.

Scarborough and Dobrich (1994) conducted a review of more than three decades of research related to the influence of parent-preschooler reading on the development of language and literacy skill. The researchers found that there is a relationship between parent-preschooler reading and preschool reading achievement. The average magnitude of the relationship is modest at best (approximately 8% by their estimate) and the variability of correlational results

from sample to sample is considerable. Though not robust, the findings of this review indicate that parents' shared reading practices are related to child literacy skills.

In response to the modest findings reported by Scarborough and Dobrich (1994), Lonigan (1994) offers reasons to be more optimistic concerning the effects of reading to preschoolers. Lonigan criticizes methodological issues, arguing Scarborough and Dobrich provide equal weight to good and poor studies and use a relatively conservative approach in their estimates. Using an alternative method of estimating an effect size from the same data, Lonigan suggests that the direct effect of preschool exposure to print can account for approximately 12 to 13% of the variance in reading achievement and emergent literacy skills. Moreover, Lonigan claims that the researchers do not account for the indirect effects that reading to preschoolers can have. Beyond reading achievement, reading to preschoolers is highly related to emergent literacy and language skills, and both of these factors operate simultaneously to influence reading achievement. With the likelihood that multiple indirect effects operate on reading achievement through parent-preschooler reading, Lonigan estimates that "preschool exposure to print accounts for more than about 16 to 20% of the variability in young school-aged children's reading skills" (p. 319). Importantly, Lonigan points out that even if young children's literacy skills are only modestly causally attributed to being read to as preschoolers, as Scarborough and Dobrich argue, small initial differences among children can be significantly magnified over time. Stanovich (1992) argues that early acquisition of better reading skills likely results in more self print exposure, which in turn can foster the development of cognitive, linguistic, and literacy abilities that provide a foundation for more reading achievement.

Other, similar research on the topic of preschool exposure to print resulted in greater effect sizes than those found by Scarborough and Dobrich (1994). A meta-analysis examining parent-preschooler reading and language abilities conducted by Bus, van Ijzendoorn, and Pellegrini (1995) revealed a much higher effect size (.59), leading the researchers to conclude that the results provided "a clear and affirmative answer to the question of whether storybook reading is one of the most important activities for developing the knowledge required for eventual success in reading" (p. 19).

Enhancing storybook reading

A meta-analysis by Mol, Bus, de Jong, and Smeets (2008) tested whether variations in parental reading affected children's language development, specifically focusing on vocabulary. A series of 16 studies involving children aged two to six years, excluding intervention studies involving teachers-and/or stranger-child book reading, suggests that enhancing the dialogue between parent and child during reading sessions strengthens the effects of book reading. When focusing on measures of expressive vocabulary, Cohen's *d* was .59. For studies that measured receptive vocabulary growth, there was a smaller effect size (*d*=.22).

There are ways to enhance typical storybook reading to further facilitate children's acquisition of important knowledge that leads to success in reading. For one, children learn more when adults read to them in an interactive manner. De Temple and Snow (2003) define the construct of 'non-immediate' talk during book reading as talk that is produced by the adult or child which goes beyond the information contained in text or illustrations. Often, non-immediate talk can be used to make predictions, to make connections to the child's past experiences or other books, to draw inferences, analyze information, to discuss the meaning of

words, or to teach a child how to tell a memory story. In a study by Snow (1991), mothers' use of non-immediate talk was positively related to their pre-school children's later performance on measures of vocabulary, story comprehension, definitions, and emergent literacy. Thus, adults can enhance children's learning during storybook reading by actively encouraging children to participate (Senechal & LeFevre, 2001). Encouraging talk about a picture or the text could enhance a child's ability to use these words at some other time.

Similarly, Dickinson and Smith (1994) examined teacher interaction styles during shared-reading on the vocabulary and comprehension abilities of 25 four-year-old children, all from different preschool classrooms. The proportion of teacher and child talk during reading that included the analysis of characters or events, predictions of coming events, and discussion of vocabulary was significantly correlated with levels of children's vocabulary and story comprehension. This study offers interesting insight into the effects of teachers' book readings on low-income, preschool children's vocabulary and story comprehension. The researchers characterized different approaches that teachers tend to use when reading picture books to their preschool class. They found that the *performance-oriented* approach, characterized by little talk during the book reading and a lot of talk before and after the reading, proved to be most effective for fostering vocabulary growth and story comprehension. When compared with styles of limited talk throughout (the *didactic-interactional* approach), or talk only during reading and none before or after (the *co-constructive* approach), the extended book introductions and talk that reconstructed the story after the reading typical of the performance-oriented approach demonstrated significant gains after one year of exposure. This study highlighted the importance of including analytical talk in book readings in order to foster literacy growth among preschoolers.

A study by Senechal (1995) further illustrated that typical storybook reading can be enhanced with some slight variations. The study included three storybook reading conditions: single-reading, repeated-reading, and questioning. In the repeated-reading and questioning conditions, the storybook was read three times – twice on one day, and a third time on the next day. In addition, the questioning condition group was asked questions about the storybook, including prompts to label pictures with words introduced in the stories. The results show that repeated readings as well as questioning during reading are beneficial for both receptive and expressive vocabulary knowledge. The researchers posited that receptive vocabulary was enhanced by listening to multiple renditions of a book, which provided opportunities to encode, associate, and store new words. Expressive vocabulary was enhanced through active responding. Planned comparisons revealed that the questioning group had significantly higher expressive vocabulary scores than the repeated-reading group, leading the researchers to conclude that asking questions is more beneficial to expressive vocabulary than to receptive vocabulary.

Reading the same stories several times is a simple strategy that can work fairly effectively in terms of enhancing vocabulary growth. In a study by Biemiller and Boote (2007), kindergarten children heard the same story read either twice or four times. The adult reader explained target book vocabulary to children. When tested on the word meanings of target book vocabulary, children gained 23% in their knowledge of word meanings when stories were read four times but only 16% when stories were read twice. There was also a clear benefit to hearing stories several times even when meanings were not taught explicitly, although more words were learned when children were explicitly instructed.

Dialogic reading

Grover Whitehurst has been instrumental in defining and assessing the concept of dialogic reading, a form of non-immediate talk related to shared-book reading. Whitehurst et al. (1994) explained that dialogic reading differs substantially from the way adults typically read picture books with children. In dialogic reading, the child becomes the storyteller: "dialogic reading involves families reading *with* their children rather than *to* their children" (Fielding-Barnsley & Purdie, 2003, p. 77). With dialogic reading strategies, the adult takes on the role of active listener, asking open-ended questions, adding information, discussing vocabulary, and asking for clarification or an increase in description of the material. Dialogic reading provides richer semantic contexts for novel words and gives children a denser exposure to book vocabulary (Temple & Snow, 2003). It also provides children with opportunities to express themselves and to build their expressive language skills with the aid of structured questions.

Arnold, Lonigan, Whitehurst, and Epstein (1994) offer a useful list of the principles of dialogic reading, which include: asking "what" questions, following answers with questions, repeating and expanding on what the child says, helping the child as needed, praising and encouraging, shadowing the child's interest, asking open-ended questions, and having fun. According to Whitehurst et al. (1994), the principles underlying dialogic reading suggest that children would benefit from active responding to picture books in which an adult is probing for expansions and asking open-ended questions that allow children to express themselves with longer utterances. Whitehurst and Lonigan (1998) posited that "dialogic reading targets expressive language skills through the use of probing, practice, teaching, feedback, and repetition" (pg. 282). Reading groups, then, should be as small as possible, with the ideal situation being one child reading with one adult. Group reading interactions may not be sufficient to produce significant improvements in children's oral language skills. Children are

thought to develop skills most rapidly when interaction occurs at a level slightly more advanced than current skills (Whitehurst & Lonigan, 1998), as reflected in the concept of scaffolding. These types of interactions are ideal in a one-on-one format. Individual children involved in a group dialogic reading format receive less opportunity for participation than children in a one-on-one format. Essentially, "dialogic reading is based on the premise that oral language is a complex skill that requires constant practice and feedback" (Whitehurst & Lonigan, 1998, p. 283).

Dialogic reading interventions

Reading interventions that attempt to enhance children's literacy and reading skills by using dialogic reading have consistently proven effective. A wide body of research indicates that children acquire specific new vocabulary in the context of dialogic book reading. The following section outlines different types of dialogic reading interventions and the respective effects of each. Despite differences in administrators (parents, teachers, volunteer tutors) and length of intervention, dialogic reading studies typically prove to be effective in enhancing vocabulary knowledge. Refer to Table 1 for a summary of the authors, methods, and results of the following intervention studies.

Hargrave and Senechal (2000) examined the effects of dialogic storybook reading on the acquisition of vocabulary for preschool children. The researchers included a regular-reading condition in which preschool teachers read in their customary manner as a comparison with a dialogic-reading condition, with teachers trained on the skills of dialogic reading. Each book was read twice, as it has been suggested that children benefit from repeated exposures to the same books (e.g. Shany & Biemiller, 1995; Senechal, 1995). After this four-week intervention,

the researchers found that the groups did not differ on the receptive vocabulary measure (Peabody Picture Vocabulary Test – Revised; Dunn & Dunn, 1981). However, the groups differed significantly on expressive vocabulary and the Expressive Book Vocabulary Test, a test of new words introduced in the ten books that children read in the daycares, which required children to label 18 nouns from pictures in the books. Despite the fact that dialogic reading did not exist in the ideal one-on-one situation, this study demonstrates that a relatively brief intervention of dialogic reading with teachers can have modest effects on the development of expressive language.

Whitehurst, Falco, Lonigan et al. (1988) conducted a one-month, home-based intervention designed to increase parental skill when reading picture books to three-year-old children by training parents with the features of dialogic reading. Compared to a control group of children whose mothers did not receive instruction but read to their children for the same amount of time, children in the intervention group scored significantly higher on a posttest measure of expressive vocabulary, and differed on receptive vocabulary, though this difference was not statistically significant. A follow-up nine months after the completion of treatment indicated that differences remained between the two groups, although they were no longer statistically significant.

On a similar vein, Arnold, Lonigan, Whitehurst, and Epstein (1994) tested the effectiveness of dialogic reading in a four-week intervention which trained a treatment group of mothers on dialogic reading, with either a videotape format or through direct interaction. When compared to a control group that did not receive any training, children with mothers exposed to dialogic reading training experienced significant gains on both receptive vocabulary and

expressive vocabulary. There were no differences between formats of training, which indicates that dialogic reading instruction by video can be as effective as direct instruction.

Whitehurst, Arnold, et al. (1994) ran a six-week dialogic reading intervention study with 73 three-year-olds from low-income families. Participants were randomly assigned to one of three conditions: school reading, school plus home reading, and an activity and attention control. For the school reading condition, teachers were trained to read to children in a dialogic reading style using a videotape training method. For about ten minutes each school day, teachers read to students in groups of no more than five children at a time. Children in the school plus home reading condition experienced small-group dialogic reading in school under the same conditions as the children in the school reading condition. In addition, a parent or primary caregiver of each child was trained to use dialogic reading at home with the same videotape procedure. Parents were encouraged to read to their child daily, though this behaviour was not tracked by the researchers. Children in the control condition engaged in play activities in small groups of five children or less under the supervision of a teacher. At post-test, children in the school condition and the school plus home reading condition had significant gains on a measure of expressive vocabulary (One Word; Gardner, 1981), while children in the control condition did not. Comparing experimental groups, students in the school plus home reading condition also faired significantly better than the students in the school condition only, indicating that increases in the amount of dialogic reading can lead to increases in expressive vocabulary knowledge. Being exposed to more hours of dialogic reading led to greater gains in expressive vocabulary. Measures of receptive vocabulary (PPVT-III; Dunn & Dunn, 1981) did not significantly differ between any of the three conditions.

A study by Whitehurst and Lonigan (1998) attempted to replicate the findings of Whitehurst, Arnold, et al. (1994) and specifically addressed the relative effectiveness of parents versus teachers in implementing the dialogic reading program with low-income children. To do this, the researchers added a third intervention group that involved only parent reading. One hundred and fourteen children aged three and four were pre-tested on three standardized tests of oral language measuring receptive vocabulary, expressive vocabulary, and verbal fluency in describing common objects. Children were randomly assigned within classroom to one of four experimental conditions. The four conditions were school reading, home reading, school plus home reading, and a no-treatment control. Both teachers and parents were trained to read with children in a dialogic reading style through videotape training. The experimental conditions lasted six weeks, at which point the children were again administered the standardized tests of oral language. A subset of 66 children provided verbal expressions during a semi-structured reading interaction. The children who were administered this assessment looked at two picture books. A familiar, adult male asked the child open-ended questions about one of the books (e.g., "Tell me about this page"). The first five minutes of children's verbalizations were scored for overall complexity, total speech production and diversity, and categories of semantic diversity (different nouns, verbs, adjectives/modifiers).

Using children's scores on the same pre-test variables as covariates, Lonigan and Whitehurst (1998) found significant differences on the measures of expressive vocabulary and verbal expression, but not on the measure of receptive vocabulary. For the expressive vocabulary measure, all three treatment groups scored significantly higher when compared with the no-treatment control group. In analyzing verbal expression, the overall MANCOVA of children's verbal productions revealed a significant effect for group. Planned comparisons

revealed that the combined home plus school reading intervention produced longer utterances, produced more words overall, produced a higher diversity of words, and produced more adjectives/modifiers than the control group. This study demonstrated that both child care teachers and parents can produce significant positive changes in the development of oral language skills of low-income children using a brief dialogic reading intervention – the average child with the highest frequency of reading in this study would have only been exposed to between 3.5 and five hours of centre-based dialogic reading. Importantly, effects were present on both standardized measures of oral language and on contextualized, spontaneous speech samples.

Using studies by Whitehurst, Arnold, et al. (1994) as a model, Fielding-Barnsley and Purdie (2003) ran an eight-week dialogic reading intervention for children over the summer in the year prior to formal schooling. Parents were trained on the aspects of dialogic reading and asked to read eight different books at least five times over the course of the intervention. Results revealed that at Time 1 (three weeks into their first year at school), the experimental group scored significantly higher than the control group on receptive vocabulary, concepts of print, measures of phonological awareness (initial consonant and rhyme), and word reading. Benefits from the intervention were still evident at the end of that same school year: the experimental group maintained a significant advantage on the concepts of print measure and the word reading task, and a moderate, though not significant, advantage on measures of phonological awareness and vocabulary. However, the design of this study leads to the question of Hawthorne effects, as the control group used for statistical analysis did not receive any type of intervention.

Invernizzi, Rosemary, Juel, and Richards (1997) analyzed the effects of Book Buddies, a one-to-one community volunteer tutorial in Charlottesville. Despite the effort of volunteers,

the researchers noted that their tutoring had not yet appeared to garner as significant gains as those obtained by trained professionals. In this study, however, non-professional volunteers were trained and supported by certified teachers and received continuous on-site training and supervision. Children in first grade received two 45-minute, one-on-one tutorials each week. Each week included repeated reading of familiar text, a phonics lesson, a writing assignment, and the introduction of a new book using dialogic reading strategies. The researchers found that alphabetic knowledge, rime and onset knowledge, and word recognition all increased significantly over a period of five months. A major limitation of this study is the absence of a control group. However, when comparing children with high attendance (over forty sessions) to those with low attendance (under 40 sessions), word recognition differed significantly between the two groups, but alphabetic and rime and onset knowledge did not. This comparison, however, does lead to the question of whether there were initial differences that between children who attended more sessions and those that attended fewer sessions.

Oral language skills and dialogic reading

Difficulty with oral language development is predictive of reading difficulty (Miller et al., 2006). Literature suggests that oral language skills contribute to academic success (Zevenbergen, Whitehurst, & Zevenbergen, 2003). Children who enter kindergarten with better oral language skills may have an educational advantage over children with lesser developed skills, as suggested by correlational studies that provide evidence of a relationship between children's preschool oral language skills and later academic achievement (Roth, Speece, & Cooper, 2002).

Given that low-income children are often at a disadvantage in oral language skills (Dickinson, McCabe, Anastasopoulos, Peisner-Feinberg, & Poe, 2003; Spira et al., 2005),

including oral language training in early interventions with at risk children may add to an early reading intervention program's success. One of the most effective ways to develop oral language skills is through dialogic shared reading (Armbruster, Lehr, & Osborn, 2000; Hargrave & Senechal, 2000; Lonigan, Anthony, Bloomfield, Dyer, & Samwel, 1999; Whitehurst & Lonigan, 2001). Dialogic reading interventions have been shown to significantly increase the number of utterances and mean length of utterance (MLU) in two and three year olds (Huebner & Meltzoff, 2005) and word production and word diversity in three to five year olds (Zevenbergen, Whitehurst, & Zevenbergen, 2003). Miller (2006) developed an objective way of measuring the quality of oral language narratives. According to Miller et al. (2006), oral language can be assessed by performance measures that characterize the productivity, fluency, lexical diversity, and narrative structure of spontaneously generated speech samples. Productivity refers to the total number of words produced and the total time taken. Verbal fluency, calculated by dividing the total number of words in the transcript by duration of the transcript in minutes (i.e. words per minute), reflects a child's general proficiency in the language. Children's vocabulary knowledge can be assessed by calculating lexical diversity, which is the number of different words produced and the number of different verbs produced (Miller et al., 2006). Narrative structure can be assessed according to the Narrative Scoring Scheme (NSS), described in further details in the Methods section, and available on the SALT website (www.saltsoftware.com).

Narrative structure is recognized as central to the development of oral and written communication skills (Miller et al., 2006). Zevenbergen et al. (2003) examined the effects of a shared-reading intervention on children's narrative skills, specifically evaluative information – the ability to use explicit references about a character's frame of mind or emotional state ("He

was sad"). One hundred and twenty-three children aged four (58% experimental group) enrolled in Head Start participated in the study. A 30-week shared-reading program that taught teachers and parents to use the technique of dialogic reading was conducted at school and at home. Children in the experimental group were exposed to shared-reading at least twice a week at school, and parents were encouraged to read the books with their children at least three times per week. Children assigned to the control condition participated in the regular Head Start curriculum. To assess narrative skill before (Time 1) and after (Time 2) the intervention, researchers read children a story about a bus and showed them a series of twelve pictures. Children were asked to retell the story while looking at the pictures. The control and intervention groups did not differ on evaluative information at Time 1. After controlling for differences in expressive vocabulary at Time 2, the intervention group included significantly more evaluative devices, references to internal states of characters, and dialogue in their narratives than the control group. This study importantly demonstrated that children who participated in a shared-reading intervention program appeared to have gained specific narrative skills, likely because they internalized the notions of story dialogue and internal states of characters (Zevenbergen et al., 2003).

Dialogic reading summary

Taken as a whole, the series of dialogic reading studies demonstrates that storybook reading which encourages more analysis from children can produce large effects on children's language, especially expressive vocabulary knowledge and oral language skill. Looking at the meta-analysis done by Mol et al. (2008), the researchers suggested that dialogic reading with older children (four to six years) does not have as great an impact as dialogic reading with younger age groups (two to three years). Older children benefited only slightly from these

interventions (d=.14) while younger children benefited moderately (d=.50). The authors posit that perhaps parents were not adept at implementing more challenging dialogue required to enhance vocabulary, or that reading sessions become less dialogic with age as children prefer to read stories with fewer interruptions. Furthermore, Mol et al. (2008) investigated differences between children who are at risk for reading difficulties versus children not at risk. Families categorized as at risk either received government support, had low incomes, or had less educated mothers. The effects of dialogic reading significantly differed between these groups, with moderate effect sizes for children not at risk (d=.53) and small effect sizes for children at risk (d=.13). This meta-analysis only included dialogic interventions that took place between parent and child. It excluded any study that also, or only, involved teacher or volunteer interventions. It is probable that parents in at-risk situations may have provided poorer quality dialogic reading.

Controlling the difficulty of text

Carver and Leibert (1995) claim that students reading a certain level of text could probably improve their reading level by reading material at or above their reading level because they would encounter enough unknown words in a context that was not relatively difficult. Also, more complex knowledge structures would be encountered and gradually acquired in appropriate text. In contrast, students who read books at levels below their current reading ability are not likely to gain in reading ability because few, if any, unknown words would be encountered and complex knowledge structures would not be encountered. In the book *The Fluent Reader*, Rasinski (2003) claims that the greatest gains in reading ability will occur when the difficulty of the text is at the student's instructional level. According to Rasinski, the optimal level of text difficulty is between 90% and 95% accuracy in word recognition.

Carver and Leibert (1995) attempted to study this effect further by offering a six-week, summer reading intervention for children in grades three to five. Children were divided into two groups – one that read library books at or below their reading level for two hours per day, and a second group that read library books that were thought to be at or above their reading level, as indicated from pre-test measures of reading ability. Unfortunately, further analyses revealed that the texts in the second group were much easier than initially assessed. The study, then, examined the effect of reading library books at levels at or below one's current reading level. Analyses of the data in this study found no consistent evidence that students in a summer reading program who engaged in reading relatively *easy* library books for six weeks gained in their reading rate or vocabulary.

In contrast to the findings of Carver and Liebert (1995), an intervention study conducted by Ehri, Dreyer, Flugman, and Gross (2007) with participants who were younger (five to six years old) revealed the reading achievements of students in an intervention group was primarily explained by reading texts at a high level of accuracy – between 98% and 100%. Children read each book with a tutor and then re-read the same book independently. During independent reading, the proportion of texts read at a level with 90% to 97% accuracy was negatively correlated with positive growth in reading, whereas the proportion of texts read between 98% and 100% was strongly positively correlated with reading growth. It may be that conventional wisdom – 90% to 97% accuracy (Rasinski, 2003) – applies to text read without prior coaching. However, during independent reading, young children seem to benefit from reading text at high accuracy levels.

It seems likely that most young children, when left to their own devices, will not be able to adequately select text that is at or just above their reading level, which would provide growth

in their reading ability and vocabulary according to Carver and Liebert (1995). Thus, it is important for any voluntary reading intervention to provide some guidance in the selection of texts. As Juel (1996) reported, the key components in teaching a child to read involves, "verbal interactions, instruction, and written materials on the right level and at the right time" (p. 288).

Literature summary

Low-income children tend to be behind in their reading ability, likely because they are read to less, have a poorer home literacy environment, and are read to without enriching parental dialogue. For children from low SES families, positive literacy experiences might occur through sources outside traditional parent-child reading experiences including community-based reading programs. Exposure to dialogic reading, greater amounts of print, repeated exposure to the same storybooks, and exposure to print at the right level can lead to gains in different reading-related skills and vocabulary knowledge for children from low-income homes. The current intervention employs all of these tools – children are provided with print materials at the right level, engage in dialogic reading with a community volunteer, and repeat their exposure to the storybooks at home.

The Present Study

The *Dialogic Reading Club* is a volunteer program that promotes literacy among children. The program consists of volunteers reading to children every Sunday at one community center, and every Wednesday at a second community center, both located in downtown Toronto. Children attend only one of the two centers. This study includes children from both centers. Both centers serve a generally low-income, ethnically diverse community. The cost of books is supported by local community organizations (e.g. the Kiwanis Toronto).

The *Dialogic Reading Club* is offered to children aged two-12. The majority of participants in the program range from age three to six; this study focuses on this group of children. A variety of picture books, beginner readers, and simple non-fiction books are assigned levels based on difficulty of text, and then grouped in sets of four or five. The leveled sets range from 1 (most basic) to 16 (most complex). The books are organized into levels by the program coordinator, who examines each book for amount of text and complexity of story material, and organizes them into groups. Since there are different classification systems used by publishers, this is an important and necessary step to group books. By providing a levelled reading system where books are chosen for children that are controlled for difficulty, the program is encouraging growth in vocabulary, reading fluency, and narrative comprehension.

In the first session that a child attends, the coordinator conducts an assessment of reading ability by choosing a picture book and reading it with the child. This general assessment of word reading ability allows the coordinator to establish the level that the child would feel comfortable reading (if able to read at all), in terms of being able to read the majority of words on the page. Typically, a child reading approximately 90% of the words correctly in a book indicates that the book is suitable for their reading level. If a child is not able to read any words, he is given sets of simple books containing minimal and repetitive text for reading with a volunteer.

Following this initial assessment, and in weeks that follow, the child is given a set of books – a book bag containing four or five titles with similar levels of difficulty – that matches his level of reading ability. During program time, the child is paired with a volunteer from the community to read the set of books together. A child is not assigned a specific, regular volunteer; rather, children are paired with any volunteer that is available. Typically, children

and volunteers read all the books in the book bag. This process usually lasts between 20 and 35 minutes.

The coordinator trains each volunteer who enters the program on dialogic reading strategies by first explaining what dialogic reading entails, and then by allowing volunteers to watch her employ the strategies with a child. The coordinator will also sit in on a volunteer's first few sessions with a child in order to offer suggestions and useful criticisms.

If the child cannot read, the volunteer and child engage in a form of dialogic reading of the book, a method of reading picture books where the child initially becomes the storyteller through a "book-walk" – looking at the pictures and predicting a logical series of events. A proper book introduction by the volunteer will alert the child to the sequence of events, tapping the child's background knowledge. The volunteer, as an active listener, asks open-ended questions about the story, prompting the child to increase the complexity and sophistication of the story being told. Following the book-walk, the volunteer will read the text of the story to the child, allowing for a natural comparison to be made between the book-walk and the actual story. At the conclusion of the book, the pair will discuss what took place, and try to relate the story to other stories or events.

If the child can read, a book-walk will still take place – the volunteer will prompt the child to flip through the book and, without reading the text, predict the events of the story based on the pictures. Clearly, this is less applicable with non-fiction material or for books without adequate pictures, which represent the minority of books used in the program. However, introductions are still provided for this material. After the book-walk, the volunteer will listen to the child read the story, helping with difficult words, asking for predictions as the story

progresses, discussing new vocabulary, and asking questions that test comprehension. During this period of assisted reading, when children cannot read a word, they are simply told what it is, and encouraged to keep reading. No concerted effort is made to have the child use phonics to identify unfamiliar words. The goal is to keep the child reading and attending to the story rather than to the process of reading.

The children bring the book bag home each week and are encouraged to practice with the texts, in line with Shany and Biemiller's (1995) claim that children benefit, in terms of word recognition and text reading rate, from repeated exposures to the same books. The coordinator encourages each parent to read the books with their children at home.

When a child returns to the program the following week, they return their book bag to a volunteer who asks the child and the parent how difficult they found the books to be. If the child and parent agree that they are willing to try a more difficult level of material, they are given a book bag for the next highest level, which they read with a volunteer. Should the parent, child, or volunteer notice that the book set is too hard, the volunteer will provide them with an easier leveled set. Typically, a child remains on the same level of material for a few weeks. The coordinator keeps track of the sets of books that are being taken by the child each week. If she notices that a child has been stagnant on the same level for too long, she will personally assess the child, establishing an appropriate leveled set. The program continues like this until the child moves into the higher stages of reading.

Hypotheses

Compared to a no-treatment control group of children that do not attend the reading program, a group of children exposed to the dialogic reading intervention are expected to differ in the following ways.

(1) The intervention group will have higher scores on post-test measures of expressive and receptive vocabulary. Children exposed to more print, appropriately-leveled print, features of dialogic reading, and repeated exposure to the same storybook can increase their vocabulary knowledge. Biemiller and Slonim (2006) claim that when an unknown word is encountered in an interesting narrative, the basis for creating meaning exists. Brief explanations can be sufficient to establish the meaning of new words. In some cases, children can even construct these meanings without explanation. Dialogic reading seems to have greater effects on expressive vocabulary, since it encourages discussion. Greater gains are expected on expressive vocabulary. The relatively long duration of the program and the repeated exposure to storybooks should also yield differences on receptive vocabulary.

(2) The intervention group will have a higher score on a post-test measure of concepts of print. By being exposed to more print, children are acquiring code-related, emergent literacy skills such as an understanding of book concepts, reading concepts, directionality, concepts of letters and words, and punctuation marks.

(3) The intervention group will be able to read more words at post-test. Although phonics instruction and decoding strategies are not a focal point of the reading program, the leveled-readers used in the program contain a high frequency of Dolch sight words. By being exposed to more text with a volunteer who helps with difficult words, and by re-reading the

story at home with parents, it is expected that children will be able to recognize and read more words.

(4) The intervention group will produce longer, more fluent, more lexically diverse, and more coherent narratives. Dialogic reading and non-immediate book talk encourage children to talk more, which supports the development of oral expressive skills. In a language sample obtained from a communicative context, it is expected that children in the reading program will provide narratives with greater productivity, fluency, lexical diversity, and provide stories with a more coherent narrative structure.

Method

Participants

Parents or legal guardians of children who were already part of the reading program were contacted in person or over the phone by the researcher and program coordinator and invited to have their children participate in the study. Parents were told about the nature of this longitudinal study at the end of November 2008 and consent forms were given to parents in the middle of December 2008. Thirty children from the two programs in Toronto were recruited as part of the experimental group. No child attended both programs. In the initial sample, there were 18 females and 12 males. Two children (one male, one female) moved before the post-test, so data is available only for a sample size of 28 children in the experimental group. This group had a mean age of 62.96 months (SD = 11.16) at pre-test, with a range of 38 months to 76 months.

Nineteen children were recruited for the control group from three separate daycares in Kitchener, Ontario. The researcher provided information about the study to the daycare

coordinators, who relayed information to parents and collected consent forms. This initial sample contained nine males and ten females. Two children (one male, one female) were not available for the post-test because they no longer attended their respective daycares. Thus, data are available for 17 children in the control group. This group had a mean age of 54.12 months (SD = 9.58) at pre-test, with a range of 39 months to 71 months.

One child in the intervention group did not attend any reading sessions because of a conflict in scheduling. This child was moved to the control group. After examining the mean ages for the intervention and control groups, the researcher was concerned with the large mean difference in age (almost nine months) between groups. In order to establish more equivalent groups on the basis of age and schooling level, the researcher excluded all the children in the intervention group who were in grade one and likely receiving formal literacy instruction at school; nine children fit this criteria and were excluded (*M* age = 74.78, *SD* = 1.56). Thus, the final sample included 18 children in the intervention group (*M* age = 58.22, *SD* = 8.55) and 18 children in the control group (*M* age = 53.11, *SD* = 9.52). Racial makeup of the intervention sample was as follows: 11 children were Chinese (eight Mandarin), four were Caucasian, two were Somali, and one was Latin American. Racial makeup of the control sample was as follows: 16 were Caucasian, two were European Canadian.

The researcher approached volunteers from the reading program at both sites with consent forms that asked for permission to sit-in on a dialogic reading session with a program participant. This was done to ensure treatment fidelity. Five volunteers agreed to participate (four females, one male). Specific demographic information was not collected on these participants.

In order to establish treatment fidelity and an understanding of how dialogic reading strategies were implemented, the researcher observed those five different volunteers engaged in a book-reading session with children. To begin, all volunteers made note of the title page, reading out the title for the child and commenting on the picture. Following this, the volunteers prompted children to take them on a 'book-walk', where children looked at the various characters and explained what was happening in the pictures. During this time, volunteers would ask open-ended questions, engaging children in a discussion on some of the more important themes in the book. For example, during a book-walk, one volunteer asked if the child noticed anything different about the main character (he was blind). The volunteer followed this question by asking the child to think about situations where being blind would be a challenge. These types of open-ended questions encourage children to respond with verbal expressions that extend beyond a few words. Since the book was focused around a central character's struggle with blindness, this discussion provided the child with a good context before actually reading the story.

Following the book walk, children were asked to begin reading the story. In all five cases observed, the child was able to read the majority of the story. During reading, the volunteer asked questions, most of which were meant to ensure adequate comprehension and were not for the purpose of engaging in extended discussion. For example, during one story, a volunteer asked three questions: "Do you know what peppermints are?; Do you know what clever means?; Who is saying that?" (referring to a quote by one of the characters). These types of clarification questions were typical of the questions posed by all volunteers.

During reading of the story, the volunteer often praised the child for demonstrating good reading skills. The volunteers helped with difficult words – in most cases, the volunteer simply

provided the child with the word when he or she was stuck. In rare cases, the volunteer broke up the word (by placing a finger over a syllable) and prompted the child to read each individual syllable (e.g. Sun-day). Overall, the majority of discussion was concentrated before reading the story. Volunteers did not typically engage in discussion after the book was completed. At the end of the reading session, every volunteer encouraged children to re-read the books at home.

Procedure

Pre-tests were scheduled for every Sunday during the month of January 2009 and the first Sunday in February 2009. Testing took place in the gymnasium of one of the community centers. Each participant was paired with a research assistant who administered the tests. Testers were colleagues of the researcher and despite having limited backgrounds in testing and literacy, they were trained thoroughly on each measure in the weeks prior to testing. During testing, the researcher was on-hand to answer questions and to ensure volunteers were correctly administering the tests.

During the pre-test, children completed a battery of tests that took approximately one-hour to complete. Most children finished in one testing session; those that did not complete the battery in one session came back the following week. The tests, described in further detail below, included measures of vocabulary, word reading, letter names and letter sounds, phonological awareness, working memory, and general cognitive ability. Upon completion of the pre-test, children received small prizes, including stickers, pencils, and erasers.

At the beginning of January 2009, the researcher visited daycare centers in Kitchener, Ontario. After explaining the nature of the study to daycare administrators (that the children would be control group participants) at eight different centers, three were interested in

participating as the control group for this study. Consent forms were distributed and collected by a supervisor at the daycare, who served as a contact for the researcher.

All children in the control group were initially tested on Thursdays or Fridays in January 2009 and in the first week of February 2009. Testers were graduate and undergraduate students in Psychology, thoroughly trained on each assessment measure. Children completed the same battery as the one completed by the intervention group. Most children finished in one testing session; others required two days of testing to complete the battery. Participants were again given small prizes for their participation.

Children in the intervention group attended the reading program in the weeks that followed their pre-test. During the intervention, the researcher visited the program frequently to ensure treatment fidelity. The researcher took notes on the majority of dialogue used by the five different volunteers while engaged in shared dialogic book reading with their children. As stated above, the researcher captured some of the dialogue that took place between child and volunteer for comparison with the standard characteristics of dialogic reading, as outlined by Whitehurst et al. (1994).

In May 2009, a *Language Questionnaire* (Appendix A) was distributed to parents in the intervention group and to the three daycare contact persons, who distributed and collected them from parents. This questionnaire was used to assess differences in home literacy environment between the intervention and control participants.

Each child was post-tested between four and five months after the pre-test; the average length of time between pre-test and post-test was 4.40 months. For the intervention group, the mean number of sessions attended was 11.11 ($SD = 4.17$). The greatest number of sessions

attended by a child was 18 and the least was five. Post-testing took place at the end of May 2009 and at the beginning of June 2009. Children in both the intervention group and the control group completed another battery of assessments which tested vocabulary knowledge, word reading, letter names and sounds, phonological awareness, working memory, general cognitive ability, oral language expression, and an activity preference questionnaire. The same testers who volunteered to help for the pre-test were provided with a refresher session on previously administered measures and given an explanation of the new measures being used in the weeks prior to the post-test. Upon completion of the post-test, all children received an envelope containing small prizes, including stickers, pencils, and erasers, a certificate indicating outstanding participation, and a brand new picture book, donated to the researcher by the Children's Book Bank of Toronto.

The researcher provided the daycare contacts with a brief write-up that summarized the nature of the study and an overview of dialogic reading methods. Parents in the intervention group received the same write-up of dialogic reading methods and a brief, simple summary of some of the key results of the research. The program coordinator was provided with all results of the study, an explanation of some of the statistics used, and will receive a copy of this report upon completion.

Assessment Instruments

All measures were administered at both pre-test and post-test unless stated otherwise. Generally, the assessment battery followed a set order, although it deviated slightly when testing instruments were not immediately available.

Word reading

The Word Identification (Word ID) subtest from the Woodcock Reading Mastery Test-Revised was used as a measure of word reading ability (Woodcock, 1998). The test consisted of the participants reading a list of words that increased in length and difficulty. There are 106 words in total and the test stopped when the participant read six words incorrectly in a row. The test has an internal consistency reliability of .92. The raw score from this test is the number of words read correctly.

The Word Attack subtest of the Woodcock Reading Mastery Test-Revised (Woodcock, 1998) measured decoding ability. The child was required to read 45 pronounceable pseudowords, and the total number of correct responses was recorded. This task has an internal consistency of reliability of .91. Floor effects were common on this measure as young children have not developed adequate decoding ability to score high on this task. The raw score from this test is the number of pseudowords read correctly.

The Dolch pre-primer and primer lists were used as simpler measures of word reading. In the Dolch pre-primer, children were shown a list of 40 high frequency sight words (e.g. a, and, me, my). If children could correctly read at least half of these words, the Dolch primer was administered. This list contained 52 sight words (e.g. all, must, well, with) that are slightly more difficult than those found on the Dolch pre-primer. Together, these two lists contain words that account for between 50 and 70 percent of the words found in most picture books (Johns, 1977). The total number of words read correctly on each list was the raw score, which was used in the analyses.

Vocabulary

Receptive vocabulary was measured with the Peabody Picture Vocabulary Test – III (PPVT-III; Dunn & Dunn, 1997), Form A. The PPVT-III prompted participants to match a stimulus word, presented orally, to one of four picture drawings. The test ended when a participant incorrectly identified eight word-picture relations in a set of 12. The words covered a variety of content areas including animals, emotions, body parts, and foods. Words included nouns, verbs, and adjectives. The average Spearman-Brown reliability coefficient for the age range tested in this study is .77. Raw scores were calculated by subtracting the number of items answered incorrectly (since the basal item) from the number of the ceiling item.

Expressive vocabulary was measured with the Expressive Vocabulary Test (EVT; Williams, 1997), a measure of expressive vocabulary and word-retrieval. This test prompted students to provide one synonym for each word provided by the tester (e.g. *dish = bowl, plate,* or *saucer*). Words were presented along with a corresponding picture that acted as a hint for participants. There were 152 synonym items. The test ended when a child failed to provide an adequate synonym for five straight test items. The EVT has high split-half reliabilities, ranging from .83 to .97 with a median of .91. Alphas range from .90 to .98 with a median of .95. The raw score for this measure is the number of correct test items passed minus the number of test items failed.

As another measure of expressive vocabulary, children's narratives (described below in the *story task*) were scored for the production of novel words. The number of different words spoken by the child during created narratives can serve as a strong, contextualized measure of oral expressive language ability (Miller et al., 2006). Miller demonstrated that vocabulary diversity, as measured in this manner, significantly correlates with age, $r = .71$ (Miller, 1987 in Miller, Heilmann, Nockerts et al. 2006).

Letter Names and Sounds

Children were shown the letters of the alphabet in random order on cue cards. First, children were asked to provide the names of each letter. Prereaders' knowledge of letter names is the strongest predictor of success in early reading achievement (Adams, 1990, p. 55). Raw scores were the number of letter names correctly identified. Following this, children were asked to provide the sound that the letter makes. Raw scores were the number of letter sounds corrected identified.

Concepts of Print

The Concepts of Print Test (Clay, 1979) is a measure of a child's exposure to books. It is divided into sub-sections that test knowledge about book concepts (e.g. where is the back of the book?), reading concepts (e.g. where do I start reading?), directionality concepts (e.g. where to go next at the end of the line), concepts of letter and word (e.g. show me the first letter in this word), and common punctuation marks (e.g. a question mark). The tester asked the relevant questions while the child viewed and pointed to a picture book. The test is scored out of 22, and the total number of items answered correctly is the child's raw score.

General Cognitive Ability

The Block Design subtest of the Wechsler Preschool and Primary Scale of Intelligence (WPPSI; Wechsler, 1991) was used to assess general cognitive ability. The test contained 14 items that required a child to use one- or two-colour blocks to re-create a specific design within a specified time limit. For test items one through eight, the tester constructed a set pattern using two to four blocks, and the participant was given thirty seconds to create the same pattern. If the first trial was not passed, the child was given a second chance. For test items nine through

14, children looked at a picture in a stimulus book and attempted to recreate the pattern with the blocks. Higher scores are given for patterns that are completed more quickly, with a maximum score of 42. Wechsler (1991) reported average reliability coefficients (odd-even correlations corrected by the Spearman-Brown formula) of .85. The test stopped when participants failed both trials for three straight test items. As a control variable, this test was only completed during the post-test.

Working Memory

The nonword repetition task from the Comprehensive Test of Phonological Processing (CTOPP; Wagner, Torgesen, Rashotte, 1999) was used to measure phonological loop storage capacity. Testers read a nonword to participants and asked that they repeat the word exactly. As the students progressed through the list of 18 pseudowords, the pseudowords increased in length and complexity (e.g., *jup, burloogugendaplo*). Raw scores are the total number of pseudowords repeated correctly. As a control variable, this test was only administered at pre-test.

An experimental task called short-list word repetition required participants to repeat two real words. The tester said a pair of two words and the participant repeated the pair (e.g. *cat-box, Cinderella-miserably*). There were eight word-pairs in total. As the test progressed, the total number of syllables in the pairs increased, challenging participants' storage capacity. As a control variable, this test was only administered at pre-test.

As another measure of working memory, participants completed the backward digit span subtest of the Wechsler Intelligence Scale for Children-III (Wechsler, 1991) at pre-test. This test required participants to listen to the tester say a string of digits (5-7-4) and then repeat the

string backwards (4-7-5). This test proved to be too difficult for children at this age and was not used as part of any analyses due to floor effects.

Processing Speed

The Rapid Automatized Naming (RAN) subtest (number naming) of the CTOPP was used as a measure of processing speed at pre-test. The task involved participants reading a list of numbers as quickly and accurately as possible. This task proved to be too difficult for some children, who did not yet know all the numbers one through nine. Because of this variability, this measure was not used as part of any analyses.

Phonological Awareness

Performance on phoneme manipulation tasks has yielded strong correlations with, or predictions of reading achievement (Adams, 1990). Phoneme oddity tasks measure participants' ability to compare and contrast similarities and differences. Bradley and Bryant (1983) found a highly significant relation between children's oddity test scores and their later reading achievement. Two oddity tasks were used in the battery: rhyme detection and phoneme detection (Gottardo, 2002). The Monster task possesses the benefit of providing culturally neutral pictures that serve to engage young children's interest in the task. In the rhyme detection oddity task, children were shown pictures of three creatures. Each creature had an irregular name, two of which rhymed (e.g. *fap, dap, smar*). Participants had to select the creature that did not rhyme with the others by pointing. In the phoneme oddity task, children were again presented with pictures of three creatures with silly names. Two of the creatures had the same starting sound (e.g. *bap, bep, gonk*). Participants had to select the creature with the name that

started with a different starting sound by pointing. Both oddity tasks contained 15 test items. The entire task was completed by each participant, regardless of performance.

The Rosner Auditory Analysis Test (AAT; Rosner & Simon, 1971) is a test that measures participants' ability to manipulate phonemes in words. The test is divided into three sections, with each section containing ten items, for a maximum raw of 30. The test stopped when children made five consecutive errors. This test involved removing parts of compound and regular words (e.g. say *pancake* without saying *cake;* say *bus* without saying */b/*). Despite stressing the importance of standardization, the researcher noticed differences in administration among testers, with some testers using fingers to represent sounds, which gave some participants in the intervention group a favorable advantage. This test was not used as part of the statistical analysis because of this discrepancy in favour of the intervention group.

The Sound Blending subtest of the Woodcock Proficiency Battery was used in both the pre-test and post-test as another measure of phonological awareness. This task involved a tester saying two separate sounds, prompting the child to put them together to form a word. However, the test is typically administered through audiotape and this was not available at the time of testing. The researcher has concerns about the reliability of this measure as testers varied in their administration of this task. For this reason, this measure was not used as part of the statistical analysis.

Motivation to read

The Activity Preference Questionnaire (Stanovich & Cunningham, 1997) was administered at post-test to garner a measure of participants' interest in reading. The task required children to select one choice out of two options (e.g. *I would rather watch television or*

read a book). Children's preference for reading over other activities has been positively correlated with measures of oral language proficiency and children's preference for watching television over other activities has been negatively correlated with measures of oral language proficiency (Stanovich & Cunningham, 1997).

Story Telling

During the post-test, children were shown the wordless picture book titled, *Frog, Where are you?* (Mayer, 1971). The tester introduced the task by telling children the name of the story, how the book was different because it did not contain any words, and that they were free to make up whatever story they pleased. The tester first flipped through the entire book, and the child was able to note the sequence of the pictures. Miller et al. (2006) explain that the story retell task requires the child to observe a wordless picture book's picture sequence, select the words, construct the sentences, and organize the sequence of propositions to retell the story as a coherent narrative.

During story-telling, the tester (always the researcher) sat near the child (but did not use pointing gestures) and provided backchannel responses only (e.g. "Oh," "Really"). If the tester realized that the child was struggling during the story and not communicating, the tester asked a probing question, such as, "what's happening now?", or "what's happening on this page?" in order to encourage talk, as is consistent with Miller et al. (2006).

Original audiotapes were transcribed by the researcher and scored for productivity, fluency, and lexical diversity. The transcripts were given to two independent coders for scoring on the narrative structure dimension. Each coder was provided with a copy of the Narrative

Scoring Scheme and examples of scored stories for Mayer's book, along with explanations, which are also available on the SALT website (www.saltsoftware.com).

The Narrative Scoring Scheme uses a 0-5 point scale for each of the following seven categories: introduction, character development, mental states, referencing, conflict/resolution, cohesion, and conclusion. Five points are given for "proficient" use, three points for "emerging" or "inconsistent" use and one point for "immature" or "minimal" use. Scores of two and four are used for intermediate performance. Scores of zero are given for poor performance, not completing the task, abandoned utterances, and unintelligibility. The scores for each characteristic were combined into a total composite score (highest possible score being 35) (SALT Software). See Appendix B for detailed scoring information and an example story with scoring. The raters assessed narrative structure on transcripts with no information on control versus intervention groups (blind coding). Participant identification numbers were used for each transcribed story, without any other identifying information.

Since narratives require a subjective score by the two independent raters, it is necessary to calculate an intra-class correlation coefficient (ICC). The ICC can range between 0 and 1; it is higher when there is little variation between the scores given to the narratives by the rater. For this study, the ICC is .96, indicating very high agreement between raters on the scores of narratives.

Questionnaire

In order to gather parental demographic information, parental language proficiency, and child activity information related to reading, a parent of each child in the study was asked to fill out a questionnaire (see Appendix A) in January 2009. Parents were asked about their native

country and language, education level, and occupation as part of the demographic profile. As outlined in the questionnaire, parents were asked to provide the same information for another adult with whom the child lives (e.g. another parent of a step-parent) or with whom the child has regular contact (e.g. a parent no longer living in the household). Parents rated their ability (and estimated the second adult's ability) to understand, speak, read, and write English on a scale 1 (no ability) to 10 (very fluent). A composite parent English language score was formed with this information, with each parent's total score (maximum 40; 10 points for each ability) added together to form a parental composite score (maximum 80; 40 for mother plus 40 for father). For one adult that did not provide information for a second adult, the score out of 40 was doubled to get a score out of 80.

For questions related to home literacy environment, the researcher used questions from past research (Hood et al., 2008; Senechal, LeFevre, Thomas, & Daley, 1998) that have correlated with various reading skills. Questions included: the estimated number of children's books available in the home, frequency of storybook reading in a typical week, estimated frequency of library visits, and child interest during storybook reading.

Results

Table 2 contains a list of the constructs in the study along with means and standard deviations, organized by group (intervention and control) and time of test (pre-test and post-test). Since standard scores are useful in understanding how the performance of this sample of children compares to other samples, these are also provided in Table 2 for the measures of receptive and expressive vocabulary. Raw scores from child assessments are used for analyses because most subtests did not generate standard scores. Table 3 contains demographic

information, including means and standard deviations for questionnaire items. Table 4 contains a correlation matrix of the dependent variables in the study, as well as the same variables from pre-test, in order to determine relationships among the variables of interest.

Control Variables. First, it is necessary to report on the control variables to highlight similarities and differences between the intervention and control group. For a summary of these control variables, refer to Table 3. To assess the equivalency of the two groups at pre-test on the control variables, *t*-tests were conducted for each control variable. The difference in age between the intervention and control group was not significant, $t(34) = 1.79$, $p = .08$. Using raw scores on the Block Design subtest, general cognitive ability did not differ between the groups, $t(34) = 0.88$, $p = .39$. Raw scores on nonword repetition did not differ, $t(34) = 0.79$, $p = .44$, nor did scores on short-list word repetition, $t(34) = 0.82$, $p = .42$, indicating that the groups had relatively similar phonological loop storage capacity at pre-test. The groups also did not differ significantly on the Word Attack test, $t(33) = 1.74$, $p = .09$ at pre-test. Floor effects were evident on this test, as expected. Differences at pre-test were present on knowledge of letter names, $t(34) = 3.10$, $p < .01$, and knowledge of letter sounds, $t(34) = 3.11$, $p < .01$ in favour of the intervention group. This difference in alphabetic knowledge was also significant at post-test. Phonological awareness, as measured by the rhyme and phoneme oddity tasks, were not significantly different: a t-test for rhyme oddity yielded a $t(34) = 0.18$, $p = .86$, and a t-test for phoneme oddity yielded a $t(34) = 0.58$, $p = .57$.

Activity Control Variables. In terms of activity control variables, there are some important differences worth noting based on the parent questionnaire. Refer to Table 3 for a summary of these activity control variables. Firstly, the control group had more children's books available in the household, $t(33) = 4.05$, $p < .001$, and in a typical week, parents of

children in the control group read more often with their children, $t(33) = 2.19$, $p < .05$. No significant differences were found for the amount of time children attempted to read books alone, $t(33) = 1.48$, $p = .15$, or for children's preference for preferring reading (p=.19) or watching television (p=.17) over doing other activities. Parents of children in the intervention group reported taking their child to the public library more often, $t(33) = 3.40$, p <.01.

Demographic variables. When compared with the control group, the intervention group is from a lower-income community with more heterogeneous backgrounds. Because of this, differences in certain demographic variables were found. On the questionnaire, parents were asked to report their comfort with speaking, reading, writing, and understanding English on a scale of one to ten, and asked to estimate their spouse's comfort on the same aspects. Out of a maximum score of 80, the intervention group averaged a composite parent English language score of 47.78 (SD = 21.76), while the control group averaged 78.06 (SD =5.15). A one-way analysis of variance indicated that this difference was significant in favour of the control group, $F(1, 33) = 31.22$, $p < .01$. In terms of education, parents in the control group were more educated on the whole, with 88% completing at least a college or undergraduate degree, compared to only 33% of parents in the intervention group achieving a similar educational level. Of the 18 parents surveyed in each group, 14 in the intervention group reported a native language other than English, compared with only two in the control group reporting a native language other than English. These differences highlight the ESL status of participants in the intervention group. See Table 3 for a summary of demographic information.

Dependent variables. In order to analyze each hypothesis, two types of analyses on raw scores were completed. Firstly, each hypothesis was analyzed with an ANCOVA that controls for the same pre-test variable. This helps to determine if there were differences between the

intervention group and control group at post-test after controlling for differences, significant or nonsignificant, that were present at pre-test. It is assumed that a covariate should correlate with the dependent variable. As seen in Table 4, all of the post-test dependent variables highly correlate with pre-test scores on the same measure (all correlations are above .80). For this reason, using the pre-test score as a covariate is rational. See Table 5 for a summary of all the ANCOVAs conducted, including information on the variable used as a covariate.

Secondly, paired-sample t-tests were conducted for both groups, analyzing differences in scores from the pre-test and post-test. See Table 6 for a summary of all the paired-sample t-tests. Oral language was examined separately in Hypothesis 4 with a MANCOVA.

Hypothesis 1: The groups will differ on post-test measures of expressive and receptive vocabulary.

On the PPVT-III, a measure of receptive vocabulary, the intervention group's raw scores improved from pre-test ($M = 63.44$) to post-test ($M = 69.04$). The control group followed a similar trend, improving from pre-test ($M = 58.61$) to post-test ($M = 63.33$) also. Controlling for pre-test differences on the PPVT-III, an ANCOVA assessing differences in receptive vocabulary at post-test revealed no significant differences between groups, $F(1, 33) = 0.18$, $p = .68$. Expressive vocabulary, as assessed with the Expressive Vocabulary Test, followed a similar pattern. The intervention group improved from pre-test ($M = 49.41$) to post-test ($M = 56.06$), and the control group improved correspondingly ($M = 49.94$ to $M = 54.11$). An ANCOVA controlling for the slight pre-test difference revealed a nonsignificant difference by group at post-test, $F(1, 32) = 1.28$, $p = .27$.

Paired-sample t-tests show that the intervention group did not significantly improve their scores from pre-test to post-test on receptive vocabulary measured by the PPVT-III, $t(17) = 1.85$, $p = .08$. The control group did not significantly improve their scores on receptive vocabulary either, $t(17) = 1.57$, $p = .14$. On the measure of expressive vocabulary, both the intervention group $t(16) = 3.71$, $p < .01$, and the control group, $t(17) = 3.01$, $p < .01$, significantly improved their scores from pre-test to post-test. Thus, the hypothesis that the intervention group would have higher vocabulary knowledge at post-test was not supported.

Hypothesis 2: The groups will differ on a post-test measure of concepts of print.

The intervention and control groups were relatively comparable on their pre-test scores on the Concepts of Print test, with the intervention group scoring an average of 10.67 correct items and the control group scoring an average of 8.33 correct items. At post-test, the intervention group improved ($M = 13.50$) but the control group remained relatively the same ($M = 8.89$). An ANCOVA assessing differences at post-test, while controlling for the initial scores between the groups at pre-test, revealed a significant difference, $F(1, 33) = 197.29$, $p < .001$ in favour of the intervention group, a moderate effect size of .22.

Paired-sample t-tests illustrate the growth in Concepts of Print knowledge for the intervention group from pre-test to post-test, $t(17) = 6.37$, $p < .001$. For the control group, a paired-sample t-test reveals no difference in scores from pre-test to post-test, $t(17) = .89$, $p = .38$. Thus, the hypothesis that the intervention group would have more concepts of print knowledge at post-test was supported.

Hypothesis 3: The intervention group will be able to read more words at post-test.

Word reading ability differed substantially at pre-test on the Word Identification subtest, the intervention group read an average of 12.06 words, compared to the control group who, on average, read less than one word ($M = 0.67$). At post-test, this large discrepancy remained – the intervention group read an average of 21.28 words; the control group read an average of 2.56 words. Controlling for the large pre-test differences on the Word ID subtest, an ANCOVA revealed significant differences at post-test for group, $F(1, 33)=5.40, p < .05$. Since there were significant pre-test differences between groups on knowledge of letter names and sounds, an ANCOVA was run controlling for three factors: pre-test scores on letter names, pre-test scores on letter sounds, and pre-test scores on the Word ID subtest. Use of letter names and letter sound scores as control variables also accounts for skills related to reading in the control group, who showed floor effects on pre-test reading scores. This ANCOVA with three covariates revealed significant differences on the post-test Word ID subtest, $F(1, 31) = 4.20, p < .05$ in favour of the intervention group, with a small to moderate effect size of .12.

A paired-sample t-test which analyzes differences in scores at pre-test versus post-test for the Word ID subtest revealed a significant change for the intervention group, $t(17) = 4.61, p < .001$, and no significant change for the control group, $t(17) = 1.76, p = .1$.

On another measure of word reading, the Dolch pre-primer, the intervention group improved the number of words they read from pre-test ($M = 20.39$) to post-test ($M = 25.78$). The control group also improved the number of words read from pre-test ($M = 2.78$) to post-test ($M =5.61$), albeit reading substantially fewer words. An ANCOVA controlling for initial pre-test differences in words read did not garner significant results for group, $F(33) = .66, p = .42$. However, a paired-sample t-test revealed a significant change in words read from pre-test to post-test for the intervention group, $t(17) = 2.69, p < .05$, but a nonsignificant difference in

words read from pre- to post-test for the control group, $t(17) = 1.88$, $p = .08$. Thus, the hypothesis that the intervention group would be able to read more words at post-test was supported.

Hypothesis 4: The intervention group will produce longer, more fluent, more lexically diverse, and more coherent narratives.

Unfortunately, pre-test scores do no exist for the narrative measure and therefore growth cannot be calculated. A MANOVA comparing the intervention and control groups on the productivity measures related to narrative ability – word productivity, fluency, word diversity, verb diversity, and quality – revealed no difference for group, $F(5, 27) = 0.88$, p = .51. Upon examining the univariate analyses related to narrative structure and ability, the intervention group showed advantages that were marginally significant, with p-values under 0.1 in all cases. The intervention group produced more words for their narratives when compared with the control group, $F(1, 31) = 3.40$, $p = .08$. The fluency of spoken narratives, calculated by dividing the total words by the total time taken, differed significantly by group, $F(1, 31) = 4.69$, $p < .05$. The intervention group had narratives that were more lexically diverse; uttering narratives with a greater number of different words, $F(1, 31) = 3.15$, $p = .09$, and a greater number of different verbs, $F(1,31) = 3.13$, $p = .09$. Finally, the intervention group produced better quality narratives for the frog story when compared with the control group, $F(1,31) = 2.93$, $p = .09$.

As clearly outlined in Table 4, narrative ability was highly correlated with receptive vocabulary, which is why it is rational to use pre-test score on the PPVT-III as a covariate. Since narrative productivity, fluency, lexical diversity, and structure are highly linked with oral language ability, a MANCOVA was run using pre-test scores on receptive vocabulary (PPVT-

III) as a covariate: this test revealed no significant effects for group, $F(5,26) = .61$, $p = .70$. The separate univariate tests that controlled for pre-test differences between groups on the PPVT-III illustrate that there are no significant differences in narrative productivity, $F(30) = 1.89$, $p = .18$, narrative fluency $F(30) = 3.08$, $p = .09$, word diversity, $F(30) = 1.75$, $p = .19$, verb diversity, $F(30) = 1.69$, $p = .2$, or narrative quality, $F(30) = 1.56$, $p = .22$.

Age is also linked to narrative structure and ability (Miller et al., 2006). Since the intervention group has an average age that is slightly but not significantly older than the average age of the control group, ANCOVAs with age as a covariate and narrative structure and ability as dependent variables were also run. When controlling for age, there are no differences in narratives between groups in the total number of words (p=.37), different words (p=.33) or different verbs used (p=.50), fluency (p=.26), or narrative quality (p=.29). Thus, the hypothesis was not supported. See Table 5 for a summary of these tests.

Discussion

The results of the present study revealed that low-income children exposed to a four-month dialogic reading intervention improved certain reading-related skills when compared with children who did not attend the program. Although significant gains were not achieved on all the expected constructs – notably, expressive vocabulary – the intervention group demonstrated greater growth at post-test on concepts of print, word reading, and some differences in narrative structure and ability.

Despite a large body of literature that supports the notion that dialogic reading is linked to the development of expressive vocabulary (Hargrave & Senechal, 2002; Lonigan & Whitehurst, 1998; Arnold et al., 1994), the present study failed to find significant effects on

expressive language development. This is inconsistent with many of the Whitehurst studies (Lonigan & Whitehurst, 1998; Whitehurst, Zevenbergen, et. al, 1999; Whitehurst, Arnold, et. al, 1994; Arnold et al, 1994; Whitehurst, Falco, et al., 1988) on dialogic reading. In this case, it is possible that more sensitive measures that assess the novel vocabulary introduced in the books may be necessary to appraise gains in expressive vocabulary. It is possible that the vocabulary in the books used in the program was not diverse enough to produce growth in expressive vocabulary, or that the assessment measure, the Expressive Vocabulary Test, was not sensitive enough to capture the true level of growth that was made by the intervention group. However, because another measure of expressive language based on narrative content was calculated and did not differ across groups at post-test, it is likely that this result is robust.

It is possible that the amount of dialogic reading in this program – in this study, the average child was only exposed to five to six hours of dialogic reading – was not enough to produce growth on expressive vocabulary above gains made by the control group. It is also worth noting that parents of children in the control group read to their children more often, and were more comfortable with English as indicated by parent English language composite scores. This positive home literacy environment may have produced gains for control children on expressive vocabulary that equaled gains made by the intervention group through the program.

The present study also failed to find significant effects on receptive language development for children in either the intervention or control group. This result is consistent with much of the literature (Hargrave & Senechal, 2000; Lonigan & Whitehurst, 1998; Whitehurst, Arnold, et al., 1994; Whitehurst et al., 1988), although it is inconsistent with the first hypothesis. The researcher felt the four-month duration of the program would be enough time for the intervention group to be exposed to more vocabulary in the context of dialogic

reading. Hargrave and Senechal (2000) feel that most dialogic reading studies do not find significant growth on receptive vocabulary because, for one, more sensitive assessment measures may be required. According to the National Reading Panel (2000), vocabulary growth in intervention studies is best assessed through researcher-developed measures, as these are more sensitive to gains achieved through instruction than are standardized tools.

Most dialogic reading studies use the Peabody Picture Vocabulary Test to measure receptive vocabulary and the Expressive One-Word Vocabulary Test (Gardner, 1981) or the Expressive Vocabulary Test (Williams, 1997) for expressive vocabulary. In one experimental study of an intervention using dialogic reading, Senechal (1997) used a more sensitive measure to assess vocabulary growth and found significant gains for an intervention group. Instead of using the PPVT-III as an assessment instrument, Senechal used her own target pictures and foils, which represented vocabulary that was actually present in the storybooks used in the intervention. Expressive vocabulary was tested by having children label target items pictured in the illustrations of the storybooks they read. This creative assessment format may provide the sensitivity needed to demonstrate actual gains made on vocabulary knowledge.

Recently, researchers have begun to examine the effectiveness of explicit, conspicuous teaching of word meanings to young children. Coyne, McCoach, and Kapp (2007) ran an intervention study with kindergarten children that compared extended vocabulary instruction, embedded instruction, and incidental exposure to target words during storybook reading. The extended vocabulary instruction condition was designed to teach children the meanings of words within storybook reading on a level that encouraged a greater knowledge of vocabulary depth including generic word meanings and specific meanings in context. Prior to each reading of the storybook, interventionists prompted students to pronounce the target words; students

then listened to and identified the word in the storybook; the sentence was re-read; students were provided with a definition of the word; the interventionist re-read the sentence, substituting the target word with its definition or a simpler synonym; students were asked to once again pronounce the word; finally, students engaged in activities that encouraged extensive processing, including formulating their own sentences with the word. When compared with incidental exposure (target words appeared in storybooks but were not taught or discussed) and embedded instruction (students were provided with simple definitions of target words when encountered in the story), extended vocabulary instruction resulted in far greater word learning. Incidental exposure to target words resulted in almost no appreciable word learning. This study importantly demonstrates that word learning should occur at a more extensive level of processing, as evidenced by diverse activities to teach meaning, in order to maximize gains. The Dialogic Reading Club uses a strategy similar to embedded instruction, which shows modest effects on word learning but may not occur at a intense enough level of processing to show considerable gains. This fact, combined with an assessment instrument that may not have been sensitive enough to capture modest growth, may have lead to the null findings in regards to vocabulary.

The hypothesis that children in the dialogic reading condition would significantly improve their performance on a Concepts of Print measure when compared with the control group was supported, demonstrating a moderate effect size of .22. Children in the program were exposed to print in a manner where they could plainly understand the concepts of a book's structure, reading concepts, and directionality. The book walks, that the volunteers prompted the children to undertake, focused the child's attention on the concept of a book and how it is read. This finding extended the benefits typically found in dialogic reading studies, which mainly

focus on vocabulary growth. The present study demonstrated that dialogic reading has the potential to enhance important emerging literacy skills, like concepts of print, which may be an important tool as a child moves into reading without assistance. Enhancing concepts of print through an intervention might be especially important for low income children who do not receive these experiences at home.

The largest difference between groups at pre-test was the intervention group's proficiency in being able to read words. This difference was greater in magnitude at post-test, indicating a higher rate of growth in word reading ability for the intervention group. Even after also controlling for letter name and letter sound knowledge at pre-test, the intervention group still demonstrated a significant gain on their word reading ability when compared with the control group, for a small to moderate effect size of .12. This advantage in being able to read more words likely stems, in part, from the program's emphasis on appropriately leveled print, books with a high frequency of sight words, and repeated reading of text. The program that the intervention group was exposed to did not make a concerted effort to teach letter-sound matching or any type of word decoding strategy – in most cases, emphasis was placed on attending to the story, so difficult words were provided for the child. Repeated reading of these texts along with having the story laid out clearly through the context of dialogic reading possibly enabled children to infer or remember many of the difficult words. Having print that is suitable for a child's level would clearly help children in being able to read most of the words on a page, which would lead them to infer some of the more difficult words. This theory is in line with the proposal made by Shany and Biemiller (1995), who note that word recognition accuracy follows from controlled exposure to new print vocabulary. Repeated reading builds

familiarity with context, and repeated practice with stories in a relevant context can improve word recognition skills (Shany & Biemiller, 1995).

In addition to standardized assessments, oral language was further assessed by performance measures that characterize the productivity, fluency, lexical diversity, and narrative structure of spontaneously generated speech samples. Evaluating the performance measures between groups, the intervention group demonstrated an advantage on all measures, though this was only marginally significant. Since these performance measures are related to expressive vocabulary (Miller et al., 2006), this is a positive result, demonstrating differences on a more sensitive, contextualized assessment. However, when separately controlling for initial receptive vocabulary ability and age, the difference between groups on the performance measures were reduced and no longer significant.

Despite large differences between the intervention and control groups on certain demographic and activity variables, some conclusions can be drawn about the program and about dialogic reading in general. The program seems to facilitate word reading ability and concepts of print knowledge. The fact that there were no significant differences on the standardized vocabulary measures and on the performance measures for narratives when controlling for age or receptive vocabulary is not necessarily a negative result if some important differences between groups are considered. The intervention group consists of children attending two community centers that serve generally low-income, immigrant families. This fact is exemplified in Table 3, which reports some important demographic differences between groups, notably native language, parental English language ability, parent educational level, the number of times parents read with their children, and the number of children's books available in the household. Taking all of these variables into account, the fact that the groups have similar

vocabulary knowledge and score similarly on the performance measures for narratives can be seen in a positive light. Considering the well-documented oral language deficiencies of low-income children (Adams, 1990; Whitehurst, Zevenbergen, et. al, 1999; Zevenbergen et. al, 2003), perhaps the program is helping enhance oral language ability to levels that match a more middle-class sample.

Despite the fact that there were no significant group differences on vocabulary, it is important to note that both groups did improve in their receptive and expressive vocabulary knowledge over the intervention period. Children in the intervention group can be classified as at risk based on either low-income status or having mothers with low levels of education. These children were also ESL learners who likely did not encounter much English at home. When compared with the control group, where mothers were well educated and children were not at risk, the fact that the intervention children improved from pre-test to post-test to a similar degree is remarkable. The same can be said for narrative ability: despite a significantly higher proportion of ESL status parents, the groups perform similarly on their ability to tell and structure a narrative. The Dialogic Reading Club, then, can be seen as program that works to effectively to close the achievement gap that exists between at risk children and children who are not at risk.

The efficacy of this dialogic reading program for improving some reading related skills has become clear over a period of four months. It is worth noting that the average child in this intervention attended 11 reading sessions; this means that the average child was exposed to only about five to six hours of intervention-based dialogic reading. This program is run using volunteers and with a minimal program budget. Books are usually purchased with money donated from local community agencies. Tutors volunteer their time; many do so in order to get

experience with children as preparation for careers in education. The tutor role does not require a highly trained educator – simply someone who has the ability to ask relevant, open-ended questions that encourage children to utter verbal expressions and who can read the words in the text the child is reading. Considering the minimal resources involved, it is clear that the present program, and dialogic reading in general, can be a useful and cost-effective tool for improving the word reading and literacy skills of young children.

The question of how this dialogic, book-focused reading program can match up with other programs, ones that incorporate a specific phonological awareness training component, is a question for future research. Much of the recent early reading intervention research has focused on phonological awareness. In these types of programs, children are explicitly taught that words are made up of individual sounds, and how to manipulate these sounds (National Early Literacy Panel, 2008). A review of 52 controlled studies by Ehri et al. (2001) using phonological awareness intervention programs for children in preschool, kindergarten and first grade showed that these programs are associated with significant improvements in phonological awareness, reading, and spelling.

A distinction can be made between code-focused and meaning-focused instruction (Connor, Morrison, & Underwood, 2007). Code-focused instruction consists of activities that aim to help children decode words fluently, including teaching phonological decoding and letter-sound connections. Meaning-focused instruction, on the other hand, encourages children to extract and construct meaning from text. Typically, meaning-focused instruction can include teaching comprehension strategies, discussion, reading aloud, peer reading, and repeated reading. In a study comparing code-focused and meaning-focused instruction, Connor, Morrison, and Slominski (2006) found that preschoolers with weaker emergent reading skills

demonstrated stronger letter knowledge growth in classrooms with high amounts of code-focused activities, while students with strong emergent reading skills demonstrated greater growth in classrooms with higher amounts of meaning-focused activities. Therefore, the child's reading skill interacts with instructional practices to produce differential gains.

When compared with other traditional reading readiness programs – ones that are code-focused – the long-term function of a book-focused emergent literacy intervention may have even more to do with children's attitude toward reading than on improving specific reading skills. Whitehurst, Zevenbergen, et al. (1999) posit that reading interventions that focus on phonics may have a positive effect on early literacy outcomes, providing children with skills necessary to decode words. But when the nature of reading changes from a focus on decoding words to reading for meaning, perhaps book-focused reading interventions like this one would be more important, affecting children's motivation to read and their ability to understand the story structure, which would enhance their appreciation for books that they are exposed to later on in elementary school (Whitehurst, Zevenbergen, et. al, 1999). It is possible that the most positive effects of the present intervention would pay the greatest dividends as children move through elementary school.

One of the few studies that combined phonological awareness training and dialogic reading was conducted by Whitehurst, Zevenbergen, Crone, Schultz, Velting, and Fischel (1999). The researchers coordinated an intervention for low-income children that combined dialogic reading with a phonemic awareness curriculum called Sound Foundations (Byrne & Fielding-Barnsley, 1991b). The treatment group was exposed to dialogic reading in the classroom (three to five times per week) and one-on-one reading at home with the same books that were being used in the classroom. In line with the Sound Foundations curriculum, teachers

exposed children to consonant sounds in the beginning and ending positions in words, engaging children with activities to further develop these skills (e.g. colour the picture that has the /s/ sound). Results of this study show significant positive effect sizes for Word Reading (Psychological Corporation, 1989) and pseudoword reading (Word Attack; Woodcock, 1987). Unfortunately, vocabulary and print concepts were not tracked longitudinally.

Limitations. In a study where random assignment was not possible, many questions can be raised about the equivalency of groups. The intervention group is from a low-income area in Toronto and despite the researcher's best efforts, a control group from the same area could not be recruited. However, it is reasonable to believe that the intervention group has a poorer home literacy environment: parents in the control group were more fluent with English, had more books available in the house, read more often with their children, and had higher levels of education.

There are some differences between groups in terms of demographic information. One important difference that must be highlighted is the large number of Chinese participants in the intervention group in comparison to the control group. This difference between groups could be not be controlled in the absence of random assignment, and there is no way of knowing if this somehow impacted the results of the present study. However, some important information about reading behaviour in the household was collected, and this seemed to indicate no distinct advantage for children in the intervention group except for a slight advantage in the number of public library visits.

There was a large difference between the groups at pre-test in their word reading ability. This difference is likely linked to differences in age, as the intervention group was, on average,

five months older than the control group. The researcher controlled for this difference when analyzing post-test data, by covarying scores on pre-test word reading. However, it is highly possible that having more knowledge of words leads to a faster acquisition of new words (Stanovich, 1992). This may be responsible for the differences found in word reading growth rate.

Another limitation of this study is the relatively small sample size. Only thirty-six children were part of the analysis of the study. The researcher was limited by the number of children enrolled in the program and the number of parents interested in having their children as control group participants, in the absence of a program. The small sample size reduced the power of some of the analyses. It is promising that significant differences were present at post-test with this small sample size, and encouraging that certain performance measures (i.e. narrative ability and structure) approached significance.

It is worth noting that the control group did not receive any sort of formal literacy program, which would have been valuable as a comparison. As it was, the intervention group received a four-month program and the control group did not receive any special program. However, the children in both groups attended school and were in junior or senior kindergarten. The mere discrepancy in having a specialized program offered to one group and not to another does not make for an ideal situation when comparisons are going to be made. The nature of this study and the timing made this problem unavoidable.

Conclusion

The generally positive and encouraging results, both in concepts of print knowledge and word reading, of this study are important for two reasons. Firstly, on a theoretical level, this

program has demonstrated that a brief dialogic reading intervention using books that are matched with a child's reading level can be a cost-effective way of improving the literacy skills of young children. Secondly, on an applied level, the program coordinator will use the positive results of this study as evidence necessary to apply for grants that will enable the expansion of this program to other low-income areas of Toronto, Ontario.

The long-term goal is to expand the program to several different low-SES areas and reach more children who need intervention programs in order to enhance their literacy skills. The researcher has made some recommendations to the program coordinator, including ensuring parents receive dialogic training as well and encouraging volunteers to discuss difficult vocabulary in greater depth. There could also be more of a focus on showing children how to break words into syllables, as was described for one of the observations, as long as it is not done so much that it interferes with talk relevant to the story. In the future, an expansion of this program would also make it easier to evaluate – random assignment could be put to use, and an increase in the number of participants would ensure adequate power for statistical tests, which was limited in the present study. Since reading is essential to success in our society and becoming increasingly important for social and economic advancement, an expansion of this high-quality, low-cost program to low-SES areas would greatly benefit these children, their parents, and society as whole.

References

Adams, M. J. (1990). *Learning to read: Thinking and learning about print*. Cambridge, MA: MIT Press.

Arnold, D. H., Lonigan, C. J., Whitehurst, G. J., & Epstein, J. N. (1994). Accelerating language development through picture book reading: Replication and extension to a videotape training format. *Journal of Educational Psychology, 86*, 235-243.

August, D., Carlo, M., Dressler, C., & Snow, C. (2005). The critical role of vocabulary development for English language learners. *Learning Disabilities Research & Practice, 20*, 50-57.

Bennet, N. (2005). Education for all: Literacy for life. *Education for All Global Monitoring Reports.* The United Nations Educational, Scientific, and Cultural Organization: Paris, France.

Biemiller, A., & Boote, C. An effective method for building meaning vocabulary in primary grades. *Journal of Educational Psychology, 98*, 44-62.

Bus, A., van IJzendoorn, M., & Pellegrini, A. (1995). Joint book reading makes for success in learning to read: A meta-analysis on intergenerational transmission of literacy. *Review of Educational Research, 65*, 1-21.

Cain, K. (1996). Story knowledge and comprehension skill. In C. Cornoldi & J. Oakhill (Eds.), *Reading comprehension difficulties: Processes and intervention* (pp. 167-192). Mahwah, NJ: Lawrence Erlbaum.

Carver, R. P., & Leibert, R. E. (1995). The effect of reading library books at different levels of difficulty upon gain in reading ability. *Reading Research Quarterly, 30*, 26-48.

Cipielewski, J. & Stanovich, K. E. (1992). Predicting growth in reading ability from children's exposure to print. *Journal of Experimental Child Psychology, 54*, 74-89.

Clay, M. M. (1979) *Stones – The concepts about print test*. Exeter, NH: Heinemann.

Connor, C. M., Morrison, F. J., & Slominski, L. (2006). Preschool instruction and children's literacy skill growth. *Journal of Educational Psychology, 98,* 665-689.

Connor, C. M., Morrison, F. J., & Underwood, P. S. (2007). A second chance in second grade: The independent and cumulative impact of first- and second-grade reading instruction and students' letter-word reading skill growth. *Scientific Studies of Reading, 11*, 199-233.

Coyne, M. D., McCoach, D. B., & Kapp, S. (2007). Vocabulary intervention for kindergarten students: Comparing extended instruction to embedded instruction and incidental exposure. *Learning Disability Quarterly, 30*, 74-88.

Crain-Thoreson, C., & Dale, P. S. (1992). Do early talkers become early readers? Linguistic precocity, preschool language and early reading. *Development Psychology, 28,* 421-429.

Crone, D. A., Whitehurst, G. J. (1999). Age of schooling effects on emergent literacy and early reading skills. *Journal of Educational Psychology, 91*, 604-614.

Cunningham, A. E., & Stanovich, K. E. (1991). Tracking the unique effects of print exposure in children: Associations with vocabulary, general knowledge, and spelling. *Journal of Educational Psychology, 83*, 264-274.

De Temple, J. D., & Snow, C. E. (2003) Learning words from books. In Anne van Kleeck, Steven A. Stahl, and Eurydice B. Bauer (Eds.) On Reading Books to Children. Lawrence Erlbaum Associates Inc.: Mahwah, NJ.

Dickinson, D. K., & Smith, M. W. (1994). Long-term effects of preschool teachers' book readings on low-income children's vocabulary and story comprehension. *Reading Research Quarterly, 29*, 104-122.

Dunn, L. M., & Dunn, L. M. (1981). *Peabody Picture Vocabulary Test-Revised.*. Circle Pines, MN: American Guidance Service.

Ehri, L. C., Dreyer, L. G., Flugman, B., Gross, A. (2007). *Reading rescue: An effective tutoring intervention model for language-minority students who are struggling readers in first grade. American Educational Research Journal, 44*, 414-448.

Ehri, L. C., Nunes, S. R., Willows, D. M., Schuster, B.V., Yaghoub-Zadeh, Z., & Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: Evidence from the National Reading Panel's meta-analysis. *Reading Research Quarterly, 36*, 250-287.

Elley, W. B. (1989). Vocabulary acquisition from listening to stories. *Reading Research Quarterly, 24*, 174-187.

Fielding-Barnsley, R., & Purdie, N. Early intervention in the home for children at risk of reading failure. *Support for Learning, 18*, 2003.

Gardner, M. F. (1981). *Expressive One-Word Picture Vocabulary Test*. Novato, CA: Academic Therapy Publications.

Goff, D. A., Pratt, C., & Ong, B. (2005). The relations between children's reading comprehension, working memory, language skills and components of reading decoding in a normal sample. *Reading and Writing, 18*, 583-616.

Gottardo, A. (2002). The relationship between language and reading skills in bilingual Spanish-English speakers. *The Lang Discord.* Lippincott Williams & Wilkins, Inc., 2001, 42-66.

Hall, W. S., Nagy, W. E., & Linn, R (1984). *Spoken words: Effects of situation and social group on oral word usage and frequency*. Hillsdale, NJ: Erlbaum.

Hargrave, A. C., & Senechal, M. (2000). A book reading intervention with preschool children who have limited vocabularies: The benefits of regular reading and dialogic reading. *Early Childhood Research Quarterly, 15*, 75-90.

Harris, M. M., & Smith, N. J. (1987). Literacy Assessment of chapter 1 and non-chapter 1 homes. *Reading Improvement, 24,* 137-142.

Invernizzi, M., Rosemary, C., Juel, C., & Richards, H. C. (1997). At-risk readers and community volunteers: A 3-year perspective. *Scientific Studies of Reading, 1*, 277-300.

Johns, J. L., Edmond, R. M., & Mavrogenes, N. A. (1977). The Dolch Basic Sight Vocabulary: A Replication and Validation Study. *The Elementary School Journal, 78*, 31-37.

Juel, C. (1996). What makes literacy tutoring effective? *Reading Research Quarterly, 31*, 268-289.

Lonigan, C. J. (1994). Reading to preschoolers exposed: Is the Emperor really naked? *Developmental Review, 14,* 303-323.